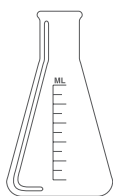


Biopharmaceutical Section



American Statistical Association

Biopharmaceutical Report

Volume 15, No. 2

Summer 2007

Chair: *Brian Wiens*

Editors: *Philip Pichotta, Richard Caplan, and Thomas Dobbins*

Note from the Editors

It is always a good feeling to have another issue of the *Biopharmaceutical Report* ready to be published. This issue's feature article describes how to use model-based drug development to improve the speed and efficiency of the drug development process. Many times we simply guess about effects but this article shows us how to do things better. Not only does the article present the methods to be used but it also provides examples to illustrate how these applications saved time and money. It is a long article but we believe that it will be worth your time to read the entire article.

We are trying to present more current information about Section activities. The highlights of the Executive Committee Meeting that was held in March in Atlanta is an example but so are the programs for the JSM and the FDA-Industry Workshop.

We also have an article about our members who were just elected as ASA Fellows and who will be honored at the JSM this summer. We congratulate those members who were elected as ASA Fellows. We hope that this article will encourage you to identify other Section members who are worthy of this honor and encourage you to start the nomination process now (the deadline for nominating Fellows is early next year).

We haven't gotten any negative feedback about the new format (single column instead of two columns) of the *Biopharmaceutical Report* so we will continue with that format.

We have several articles in the works for future issues that we hope you will be interested in. One article will discuss the new Webinar initiative to present statistical topics to our members and the response of the participants. Other articles we hope to have are on Patient Reported Outcomes and Oncology Studies. As always we are open to suggestions for additional articles. ■

Contents

Featured Article

Model-Based Drug Development – A New Paradigm for Efficient Drug Development
Kenneth G. Kowalski, Wayne Ewy, Matthew M. Hutmacher, Raymond Miller, and Sriram Krishnaswami 2

Biopharmaceutical Section News

Note from the Editors 1

Letter from the Chair
Brian Wiens 2

Biopharmaceutical Section Members Elected
ASA Fellows
Keith Soper 22

Highlights of the Biopharmaceutical Section
Executive Committee Meeting, Atlanta,
Georgia, March 12, 2007
Brian Wiens 23

Biopharmaceutical Section Program for
JSM 2007
Amit Bhattacharyya 24

2007 FDA/Industry Workshop, September
17-19, Arlington, VA
Dionne Price and Matilde Sanchez 25

Let's Hear from You 26

Letter from the Chair

Brian Wiens

As we move into the summer months, activities sponsored by the Biopharmaceutical Section are heating up right along with the weather.

The Section is involved in many meetings during the summer and early fall. The Midwest Biopharmaceutical Statistics Workshop in May is co-sponsored by the Section. By the time this issue is published, MBSW will probably be over, so you can make plans for next year.

The Section is very active at the Joint Statistical Meetings, to be held July 29 – August 2 in Salt Lake City. Amit Bhattacharyya been working for over a year and the result is an excellent program of technical sessions, roundtable discussions and short courses. Shuguang Huang is coordinating the Best Contributed Paper award and Christie Clark is chairing the Student Paper Award committee. While you are at JSM, please stop by the Section mixer and business meeting. Check the JSM program for the date, time and location. Members of the Executive Committee will be there to talk about the Section or any other topics of mutual interest, and it is a good way to meet others in the industry and make contacts.

Dionne Price and Matilde Sanchez, co-chairs of the FDA-Industry Statistics Workshop, and their organizing committee are planning the 2007 version of this popular event to be held from September 17-19 in Arlington, Virginia. Registration is scheduled to open in June.

A new initiative this year is the web-based distance learning program. Alex Dmitrienko inaugurated this event with a webinar in March, and several more have been planned. The Section subsidizes part of the cost, allowing members to participate at a very low price. This is part of the initiative of the Executive Committee of the Section to provide more services to our members. We anticipate that these webinars will continue until technology provides an even better alternative for us.

Of course none of our activities would be possible without a legion of dedicated volunteers. Check out the list of volunteers, elected officers and appointed functions on our website: <http://www.amstat.org/sections/sbiop/>. You can also find links to information about the webinar series, awards and, of course, *Biopharmaceutical Report*.

Despite all of these activities, there is still room for more. The Section has a healthy budget thanks to our members and our corporate sponsors. Executive Committee members are eager to hear from any member with an idea of how to provide more services to our members. Membership on the Executive Committee is diverse, with representation from many pharmaceutical companies, the FDA and elsewhere. Any Section member is welcome to contact any Executive Committee member to discuss opportunities for volunteering or new initiatives. With over 2200 members, there are surely a few who could benefit from your ideas, so please be in touch. ■

Model-Based Drug Development – A New Paradigm for Efficient Drug Development

**Kenneth G. Kowalski, Wayne Ewy, Matthew M. Hutmacher, Raymond Miller,
and Sriram Krishnaswami**

Pfizer Global Research and Development, Ann Arbor, MI 48105

1 Introduction and Preamble

In March of 2004, the US Food and Drug Administration (FDA) posted a white paper on their website which discusses the growing concern regarding stagnation in the drug development pipeline (FDA, 2004). In particular, the costs of drug development continue to increase while the number of new drug applications (NDAs) submitted to the FDA has declined over recent years. In the FDA's view, the problem is that the applied sciences for drug development have not kept pace with the basic sciences for drug research. Specifically, the medical advances in

drug discovery have not resulted in lower failure rates in late stage clinical development. The high failure rate in late stage clinical development continues to drive up costs for new medicines. The FDA urges that more applied scientific work be done to create new tools to fundamentally improve how safety and effectiveness of drug products can be demonstrated faster, with more certainty, and at lower costs.

One of the opportunities to modernize the critical path in drug development is to employ the concept of ‘*model-based drug development*’. The FDA characterizes this approach as the development and application of ‘*pharmaco-statistical*’ models of drug efficacy and safety from preclinical and available clinical data to improve drug development knowledge management and decision-making. This article outlines the key components of a model-based drug development approach and discusses the challenges to a successful implementation. For our purposes, we will focus on application of model-based drug development to phases 2 and 3 of clinical development; the concepts provided in this article can be extended to earlier phases including preclinical and translational (prediction from animals to humans) research.

Models are important to understand and synthesize the information we collect in clinical development. Models can be empirical, based on few assumptions, or mechanistic, based on rich assumptions leveraging our understanding of the disease process and clinical pharmacology of the drug.

The model development process that characterizes our current knowledge of a drug candidate is an iterative process, eloquently summarized by Sheiner (1997) as the ‘*learn-confirm*’ paradigm. Sheiner proposed that clinical drug development from first-in-human testing through regulatory approval can be characterized by two major learn-confirm cycles. The first cycle corresponds to phases 1 and 2A, where one learns about what dose is tolerated (phase 1) in normal subjects and confirms that this dose has promise of efficacy in a selected group of patients (phase 2A). The end of phase 2A is a decision point: If there is a sufficiently positive indication of efficacy and lack of toxicity to warrant further investment in the development of the drug, a larger and more costly learn-confirm cycle is begun corresponding to phases 2B and 3/4. In the second cycle (phase 2B), we learn how to use the drug in the target patient population, focusing on the benefit/risk trade-off. In the confirm step (phases 3 and 4), we hope to demonstrate, in a large and representative patient population, that acceptable benefit/risk is achieved, providing a basis for approval to market the drug.

Learning and confirming are distinct activities reflecting different objectives for clinical study design and analysis. It is useful to distinguish between ‘learning’ and ‘confirming’ questions. Learning questions are typically of the “how much?” variety, such as “what is the effect size?” or “what is the shape of the dose response?” or “what dose is necessary to achieve a desired effect size?” Confirming questions are typically of the “yes/no” variety, such as “is the effect size big enough?”

Models can be developed to quantify what we’ve learned. The exploratory development of assumption-rich semi-mechanistic models from the available clinical data, based on our understanding of the clinical pharmacology of the drug and the disease process, to characterize the longitudinal exposure-response relationship of the drug, is a learning activity. Evaluation of the predictive performance of these models often referred to as ‘*model validation*’, can be used to confirm what we’ve learned from independent studies not used in the development of the model. The strategic application of these semi-mechanistic models may suggest hypotheses which can be tested in a confirmatory trial, the analysis of which will often employ empirical (analysis) models, which require few assumptions. The integration of these mechanistic and empirical models, and their strategic application to clinical study designs to address both learning and confirming questions, resulting in improved drug development decision-making, are the hallmarks of this model-based drug development paradigm. We believe that this new paradigm holds promise to improve the efficiency and cost-effectiveness of clinical drug development.

We first define some basic terms that are used in this article and review the current state of drug development. *Pharmacometrics* is the quantitative science formed by the overlap of the pharmacology and statistics disciplines. The focus of pharmacometrics is the development of mathematical models that characterize the pharmacokinetics (PK) and pharmacodynamics (PD) of drugs. *Pharmacokinetics* is the study of what the body does to the drug, characterizing the absorption, distribution, metabolism, and elimination of the drug in the body. PK models

The high failure rate in late stage clinical development continues to drive up costs for new medicines.

describe the time-course of drug plasma concentrations in the population and help identify sub-populations where drug exposure may differ. *Pharmacodynamics* is the study of what the drug does to the body, characterizing the time-course and exposure-response relationships of both desirable (efficacious) and undesirable (adverse) effects. PK/PD models relate biomarkers and/or clinical measures of efficacy and safety to PK measures of exposure. *Disease progression* (or regression) models are developed to describe the time-course of disease under placebo conditions or in the absence of treatment.

An important goal of this modeling is to determine the dose(s) and regimen(s) that achieve the target clinical effect while minimizing the undesirable side effects. The strategic application of these models to make decisions about designs of future clinical trials often require the use of simulation techniques because the parameter space for these models is inherently complex and nonlinear. We collectively refer to the simulation techniques used to strategically apply these PK/PD and disease models (as well as other mechanistic and/or empirical models) to evaluate clinical study designs as *clinical trial simulations* (CTS). A cornerstone of model-based drug development is the use of sound statistical modeling principles to leverage our knowledge of the drug's pharmacology and to apply this knowledge using sound statistical design principles. As a result more informative clinical trials will be conducted that lead to improved drug development decision-making.

The current practice of drug development often emphasizes a compartmentalization of the key clinical drug development disciplines, particularly Clinical, Clinical Pharmacology, and Statistics. Each discipline tends to 'own' a piece of the development process, as depicted in Figure 1.

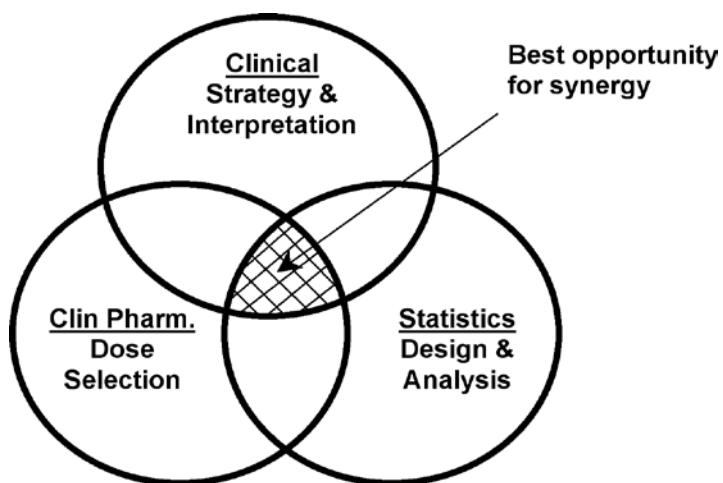


Figure 1. Current Paradigm – A Compartmental View of Clinical Drug Development

Historically, clinicians have focused on the strategy of the clinical development plan and the clinical interpretation of the trial results. Clinical pharmacologists on the other hand, view the dose selection rationale as one of their key contributions to the drug development process. Finally, statisticians have a strong sense of ownership over the design and analysis of the clinical trials. This does not suggest that the three disciplines work in isolation or don't value contributions from the other disciplines. Quite the contrary, the most effective clinical drug development teams collaborate at the intersection of these disciplines to collectively provide input on the strategy, dose selection rationale, study design, data analysis, and interpretation of the results. The use of models and clinical trial simulations to characterize what we know and what assumptions we are willing to make during the course of drug development is a useful way to integrate this cross-disciplinary knowledge base.

Presently, these modeling efforts are performed in an ad hoc and opportunistic, rather than routine, fashion, often in a crisis-management mode to address unexpected findings that threaten to delay or derail the drug development program. In this ad hoc mode it is difficult to plan and integrate these modeling activities into the operational plans for drug development. The pharmaceutical industry has optimized the mechanical process of conducting and reporting the results of individual clinical trials, but the processes that facilitate learning from these trials, such as with modeling and simulation, have largely been ignored.

We anticipate that effective clinical drug development teams will routinely collaborate at the intersection of these three disciplines building and interpreting models that summarize the current knowledgebase of the drug. The development of pharmaco-statistical models will be the cornerstone to effectively integrate, summarize, and communicate the knowledge gained from each discipline's contribution. With this approach all disciplines equally 'share' and contribute to the development and application of these models. To facilitate this model-based approach and realize its full potential, clinical drug development enterprises may need to restructure their organizations and

re-engineer processes. We elaborate on some challenges to the successful implementation of these concepts in the Discussion section below.

Model-based drug development can be described as having six major components, which are depicted in Figure 2. We discuss the key activities associated with each of these components in the following section. We then illustrate their strategic application with selected examples from Pfizer's drug development experience. We conclude with a discussion of the challenges that three scientific disciplines and senior management must address for this new paradigm to reach its full potential.

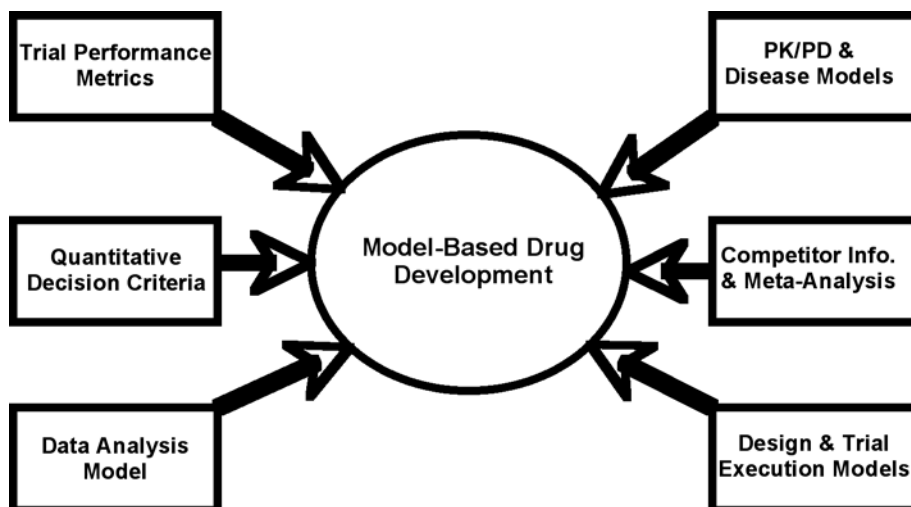


Figure 2. New Paradigm – The Six Components of Model-Based Drug Development

2 Model-Based Drug Development

The six major components of model-based drug development are briefly described in this section.

2.1 PK/PD and Disease/Placebo Models

In this section we give a high level overview of PK, PK/PD, and disease/placebo models, and their integration into population models that describe the mean and variability of the subject-level data.

PK Models. The four basic pharmacokinetic (PK) processes (absorption, distribution, metabolism, and elimination) can be described as the transfer of drug between hypothetical ‘compartments’ which can be characterized mathematically by systems of differential equations (one differential equation for each compartment). The solutions to these systems of differential equations give rise to compartmental models. The term *linear pharmacokinetics* refers to a system in which the differential equations have constant coefficients, known as proportionality constants or first-order rate constants. In such a system, the amount of drug transfer is directly proportional to the amount in the compartment. A property of linear pharmacokinetic systems is that analytic solutions can be derived from the system of differential equations (e.g., see Gibaldi and Perrier, 1982). One of the simplest models, for describing the time-course of plasma concentrations following a single oral dose of drug, is the one-compartment model with first-order absorption and elimination, for which the solution is the Bateman function, given by the expression:

$$c_p = f(D, t) = \frac{FDk_a}{V(k_a - k_e)} (e^{-k_e t} - e^{-k_a t})$$

where c_p denotes the plasma concentration as a function of dose (D) and time (t), with parameters: F, the fraction of dose absorbed; V, the volume of distribution (a scalar quantity to account for the fact that we measure concentrations rather than amounts); k_a , the first-order rate constant for absorption of drug; and k_e , the first-order rate constant for elimination of drug. Note that F and V are non-identifiable and only their ratio, V/F, the ‘apparent’ volume of distribution can be estimated. For this model, the plasma concentration is directly proportional to dose. This dose-proportionality feature holds for all linear PK models and hence, it is easy to predict the time course of exposure at other doses. In linear PK systems, it is also straightforward to predict the time course of exposure for any arbitrary dosing regimen once the parameters have been estimated from single dose data (e.g., see Gibaldi and Perrier, 1982).

Of course, for some drugs, the linear PK assumption does not hold and more complex nonlinear PK models are needed to describe the relationship to dose (i.e., lack of dose proportionality) and the accumulation of drug following multiple dose administrations. In this setting, both single and multiple dose data, over a wide range of doses, may be necessary to develop a PK model that adequately describes the nonlinear kinetic behavior.

PK/PD Models. When the pharmacodynamic (PD) response is governed by drug binding to a single receptor site, which in turn directly elicits the response, the result is a monotonic relationship with exposure, with an asymptotic maximum effect that is reached when the amount of drug at the effect site exceeds (saturates) the number of receptors available for binding. The basic model that describes this saturable binding process is the hyperbolic equation, also known as the Emax model, given by the expression:

$$E = f(c_e) = \frac{E_{\max} c_e}{EC_{50} + c_e}$$

where E denotes the effect or response (biomarker, clinical outcomes of efficacy or safety, etc.), which is a function of the concentration at the effect site (c_e), with two parameters: E_{\max} , the maximum effect the drug can achieve, and EC_{50} , the concentration at the effect site that achieves 50% of the maximum effect. The potency parameter EC_{50} can also be interpreted as the inverse of the affinity of the drug for the receptor. When two or more drugs have the same mechanism of action, the drug with the higher affinity will have a lower EC_{50} . When two or more drugs have the same mechanism of action, perhaps differing in their affinity for the receptor (EC_{50}) but subject to the same maximum effect (E_{\max}) the Emax model can be extended to describe the PD response to multiple drugs, with only the potency parameters (EC_{50}) being drug-dependent. Moreover, if a correlation has been established between potency estimates from in vitro or animal models and estimates from clinical data, the PD responses can be predicted for a new compound in the class based only on preclinical estimates of potency, together with the drug-independent parameters of the PD model.

When the effect site is in rapid equilibration with the plasma, the plasma concentration (c_p) is often used as a surrogate for the effect site concentration (c_e) in models such as the Emax model above. However, when temporal delays occur, such that the time of the peak response (E) is delayed or 'out of phase' with the time of the peak plasma concentration (c_p), a phenomenon known as hysteresis, then more complicated PK/PD models arise. For example, the hypothetical effect compartment model, proposed by Sheiner et al. (1979), assumes that hysteresis can be accounted for by the slow distribution of drug to the effect site, which can be described by incorporating an additional parameter, the equilibration rate constant k_{e0} . With this model one can infer the shape of the time-course of the effect site concentrations (c_e), up to a scalar quantity, solely from the PK and PD measurements without measurement of the effect site concentrations. The predicted effect site concentrations will be 'in phase' with the time course of the response (E) and the Emax model can be based on the predicted c_e . The solution for the effect site concentration is obtained from an additional differential equation that describes the transfer of drug to the hypothetical effect site. Holford and Sheiner (1981) provide effect compartment model solutions for a variety of linear PK compartmental models for both oral and intravenous administration. These effect compartment models contain the equilibration rate parameter (k_{e0}) in addition to the PK parameters. When distribution of drug to the effect site is rapid (as k_{e0} approaches infinity) the effect site concentration (c_e) model collapses to the solution for the plasma concentration (c_p) model. One of the key features of the effect compartment model, where the equilibration rate is governed by a first-order process, is that the degree of hysteresis, the time lag between the peak PK and PD responses, is invariant with dose. Dose-dependent hysteresis often indicates that some other process is involved.

Another class of models that can account for dose-dependent hysteresis, as well as time lags on the order of days, weeks or even months, is the indirect response model, as proposed by Dayneka, Garg, and Jusko (1993). These models assume that the effect site is in rapid equilibration with the plasma compartment and that the temporal delay in response is the result of a cascading set of events after the drug binds to the receptor. These models are more complex and the system of differential equations are nonlinear in the parameters and, hence, do not have analytical solutions. Numerical methods are required to obtain estimates of the model parameters.

Other classes of models include transduction and cell proliferation models. They are particularly useful in describing the mechanism of action of oncology drugs but are beyond the scope of this article.

Disease/Placebo Models. Disease and placebo models tend to be more empirical in nature, in large part because there may be many factors, both known and unknown, that can influence the time course of the response under placebo conditions or in the absence of treatment. Some of these factors include: the natural time course of the disease (often referred to as disease progression or regression), diurnal variation (circadian rhythm), a true placebo effect, and regression to the mean. The latter can occur when an entry criterion for inclusion in the study is that the baseline value for the characteristic of interest must exceed a threshold value. In this setting, patients may meet the inclusion criteria simply because they were measured in a transient ‘flared’ state, so that over time, even without any treatment, the patient’s response will decline naturally to the normal steady-state level.

Depending on the length of the study and the frequency of the measurements over time, the time-course of disease can be modeled using linear or exponential functions of time; or more complex functions of time, such as sine or cosine functions, to account for cyclical patterns of a circadian rhythm. When combining disease and PK/PD models, a variety of assumptions might be postulated for how the disease/placebo and drug effects interact. For instance, the apparent time of course of the disease/placebo response could be independent of the drug effect and hence an additive relationship could be postulated. Or the disease/placebo and drug effects might interact multiplicatively such that the drug effect is some fraction of the placebo response. Other, more complex relationships can also be postulated, such as having the drug influence parameters of the disease progression model.

An example of a PK/PD model with multiplicative placebo and drug effects is given by the expression:

$$E = f(t, c_p) = \left(E_o e^{-k_{plc} t} + E_{ss} \right) \left(1 - \frac{E_{max} c_p}{EC_{50} + c_p} \right)$$

where E denotes the effect or response, a function of both time and plasma concentration, which is governed by the parameters: E_o , the baseline response; k_{plc} , a first-order rate constant that describes the exponential decline of the response over time; E_{ss} , the steady-state response in the absence of treatment or on placebo; EC_{50} , the potency parameter; and E_{max} , the maximum drug effect. In this “inhibitory Emax model”, E_{max} represents the maximum fractional change in the drug effect relative to the placebo effect; if E is bounded to be ≥ 0 , then $0 < E_{max} \leq 1$.

Population Models. Population models, also referred to as pharmaco-statistical models, integrate the subject-matter PK, PK/PD or disease models with statistical models that describe the mean and variance of the responses in a population of individuals. We often assume that each subject in the population has the same structural form for the PK, PK/PD, and disease response models, differing only in some subject-specific parameter values. This is conceptually similar to random coefficient linear models. For example, each subject may have different values of PK (e.g., V/F , k_a , k_e), PK/PD (e.g., E_{max} , EC_{50}) and/or disease (e.g., E_o , k_{plc} , E_{ss}) parameters. Population models are typically defined using a two-level hierarchy consisting of within-subject and between-subject models, each of which might be non-linear. The general form for the within-subject model is given by the expression:

$$Y_i = F(\mathbf{x}_i; \boldsymbol{\theta}_i) + \boldsymbol{\varepsilon}(\mathbf{x}_i; \boldsymbol{\theta}_i, \boldsymbol{\gamma})$$

where \mathbf{Y}_i denotes a vector of individual responses for the i^{th} subject, \mathbf{F} denotes a vector-value function of model predictions (e.g., a compartmental PK model or an Emax model) that depends on \mathbf{x}_i , a vector or matrix of independent variables (e.g., time and dose), and $\boldsymbol{\theta}_i$, a vector of subject-specific parameters (e.g., V/F , E_{max} , etc.). $\boldsymbol{\varepsilon}$ is a vector of within-subject measurement errors that may depend on the independent variables, the subject-specific parameters, as well as $\boldsymbol{\gamma}$, a vector of variance parameters. The vector of within-subject measurement errors ($\boldsymbol{\varepsilon}$) is assumed to have mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$.

At the second level of the two-level hierarchy, the between-subject model is defined as:

$$\theta_i = G(\mathbf{z}_i; \boldsymbol{\theta}, \boldsymbol{\eta}_i)$$

where \mathbf{G} is vector value function that depends on a vector of covariates \mathbf{z}_i (e.g., body weight), a vector of fixed effects population (typical individual) parameters $\boldsymbol{\theta}$, and a vector of between-subject random effects $\boldsymbol{\eta}_i$ that accounts for unexplained variation between the subject-specific parameters (θ_i) and the population fixed effects parameters ($\boldsymbol{\theta}$). The vector of between-subject random effects ($\boldsymbol{\eta}_i$) is assumed to have mean vector 0 and covariance matrix $\boldsymbol{\Omega}$. For more details on nonlinear mixed effects models see Davidian and Giltinan (1995), and Vonesh and Chinchilli (1997).

Population models are used to estimate the population parameters (fixed effects) of the PK, PK/PD or disease models, to quantify the between-subject parameter variability, and to identify factors (covariates) that may influence these parameters. These population models are used to provide an integrated summary of the time-course of exposure (PK) and the exposure-response relationship (PK/PD) for both efficacy and safety endpoints in both healthy volunteers and patients. They have also been used to support dose selection, to identify subpopulations where dose adjustments may be required, and to provide support for drug label recommendations. For planning subsequent clinical trials, these models can also be used to predict the anticipated treatment effect size and as a basis for CTS data generation models.

2.2 Meta-Analytic Models of Competitor Data

The commercial opportunities for a candidate drug are highly dependent on the competitive landscape, particularly the competitor drugs' treatment effect sizes and time-to-response patterns. The systematic collection and quantification of published results (e.g., literature, summary basis of approvals, etc.) on these competitors is needed to fully understand the competitive situation. Typically, these published sources provide only study-level means and standard deviations (or standard errors). From the data for all relevant studies, a meta-analytic model can be developed to characterize the dose- or exposure-response relationship and the time-course of response, typically using nonlinear mixed effects analyses with study as the cluster variable. In this analysis, the individual group-time means within a study are weighted proportional to \sqrt{N} (or inversely proportional to the standard error of the mean) to account for the varying precision of the estimates; the treatment means are assumed to be independent across studies, while the correlation of means within a study is accounted for by study-level random effects. See Mandema et al. (2005) for a meta-analysis example characterizing the dose-response relationships for lipid-lowering agents.

There are many technical challenges with performing such longitudinal, dose-response meta-analyses, including: publication bias, incomplete description of trial design and methods, appropriately accounting for the correlation between time points within each study, and how best to incorporate patient-level data, when available for some drugs. We briefly discuss each of these challenges.

Publication bias: There is a tendency to publish positive results, while results of negative studies might not be (as promptly) published. Analyses based solely on public domain data may provide an upwardly biased estimate of the effect sizes. Recently, there have been efforts by pharmaceutical companies to more fully disclose clinical trial summary results, regardless of whether the results are positive or negative. The trade organization, PhRMA (Pharmaceutical Researchers and Manufacturers of America), has created a clinical study results database specifically for this purpose (www.clinicalstudyresults.org).

Inadequate description of patient populations and methods: To appropriately pool data across studies requires that the patient populations be similar or that any differences are adequately accounted for; unfortunately, the inclusion and exclusion criteria might not be fully described in the publications. Moreover, ascertaining the data analysis and missing data imputation methods can sometimes be difficult. The most challenging part of performing a meta-analysis is identifying a representative subset of studies for which sufficient information is available to allow their data to be appropriately pooled.

Correlation over time: The correlation structure between the group means over time is often not available as part of the meta-data. For example, in a parallel group study with repeated measures of the same individuals, the reported meta-data might be just the means and standard errors at each time point, with no information regarding the correlations over time. An analysis that assumes independence across time points within each treatment can result in standard errors for the parameter estimates being too small. The resulting confidence limits must be carefully interpreted to minimize the impact of this bias. In the situation where patient-level data is available for one or more of the drugs, from which the correlations can be estimated, it might be reasonable to assume the same correlation structure applies to the published grouped data. How best to incorporate the correlation structure from patient-level data with the study-level meta-data in longitudinal meta-analytic models is an area that requires more statistical research.

When developing longitudinal meta-analytic models combining patient-level data with meta-data from external sources, the patient-level data is commonly summarized in the same form as the external meta-data, allowing the analysis to be based on a single longitudinal model. This might be inefficient from both statistical and resource perspectives. A more efficient statistical method might be to jointly estimate both the patient-level and meta-analytic models from the external meta-data without having to perform data-reductions of the in-house data. However, more statistical research is needed to identify the best methods for such an analysis. From a resource perspective, it is also inefficient and a duplication of effort to model the in-house meta-data since it is likely that a patient-level longitudinal exposure-response model will also be developed for the in-house data. There may be differences in predictions from the patient-level and meta-analytic models especially when the structural models are different. Additional time and resources may be required to reconcile differences in patient-level and meta-analytic model predictions especially if they result in different recommendations about further development of the drug. Moreover, extrapolations with meta-analytic nonlinear models can be more problematic than for patient-level population models. This is because the shape of the time-course and dose-response relationship for the population average can be different from that of an individual's response when the true underlying response model is nonlinear. Hence, if the basic premise of the population model holds where each individual has the same structural model with subject-specific parameters, then a structural model that is developed based only on population averages (meta-data) can result in model misspecification. This model misspecification may not be important for interpolation if the meta-analytic model adequately describes the summary data over the range of the design but can lead to different predictions from a patient-level population model when extrapolating beyond the range of the data and design.

The quantitative scientists on the project team should collaborate to address these technical challenges, seeking sound statistical inference, which integrates both the internal and external information, to facilitate more informed drug development decision-making.

2.3 Design Considerations and Trial Execution Models

Pharmaceutical companies as well as regulatory agencies are increasingly advocating the use of adaptive designs to more efficiently make internal decisions and control costs of clinical studies, especially during the learning phase of a drug development program. These 'enhanced' designs focus on learning from the data as it is collected and making informed decisions earlier in the development. For instance, certain design features such as randomization ratios might be modified (in an appropriately blinded and controlled fashion) to focus enrollment on doses showing the most promise and/or to terminating doses deemed futile. Such designs can reduce trial cost as well as minimize exposure of patients to inactive (or unsafe) treatments. The design must be 'fit for purpose', so the performance to efficiently and effectively address the trial's objectives should be evaluated. CTS techniques can be employed to evaluate the performance of potential designs, including fixed and adaptive/flexible, so that the project team can fully understand the operating characteristics of each design based on the current available information about the drug before a specific design is selected. This is discussed further in Section 2.6 (Trial Performance Metrics).

During the conduct of the trial there are other factors besides the underlying PK/PD and disease properties that contribute to the overall performance of the study. Factors such as compliance and dropout can also have a substantial impact. Kastrissios and Girard (2003) define *trial execution models* as models that describe protocol deviations from the specified study design. Examples of trial execution models include models that describe the

dropout (failure to complete full duration of study medication) and/or compliance behavior (failure to strictly follow dosage regimen). We expand on the definition of Kastrissios and Girard to also include models that are employed in the conduct or execution of the design. Such an example is a model used to analyze dose-response data during an interim analysis where an adaptive dose allocation scheme is used to determine the optimum allocation of new patients to doses. Berry et al. (2002) and Smith et al. (2006) discuss application of a normal dynamic linear model (NDLM) as an adaptive dose allocation trial execution model for Bayesian adaptive designs. We advocate routine development of trial execution models, so that more quantitative assessment of protocol deviation impact can be made.

With traditional clinical trial planning, the sample size is set to achieve the desired power, at a selected significance level, assuming a specific fixed treatment effect (Δ) and variance, and perhaps inflated to account for anticipated drop-outs. In quantitative drug development, the PK/PD, disease, meta-analytic, and trial execution models are used together to predict the treatment effect (Δ) as a function of dose, regimen and time. Moreover, uncertainty in the prediction of Δ can be taken into account from the uncertainty in the parameter estimates of these models. For example, trial-to-trial variation reflecting the uncertainty in the parameters (and indirectly in Δ) can be accounted for by parametric or nonparametric bootstrapping techniques. CTS techniques are then employed using the models together with the bootstrap vectors of parameter values for each simulated trial reflecting the uncertainty, to simulate hypothetical data for each of many simulated clinical trials for each potential design under consideration. Essentially, this CTS approach facilitates the calculation of 'marginal' power averaged over the uncertainty in the prediction of Δ . This 'marginal' power calculation leads to a larger sample size relative to assuming the mean Δ is known (without uncertainty) but a smaller sample size relative to the worst case one might assume over the range of plausible values of Δ given this uncertainty.

If the primary endpoint involves an imputation method to account for dropouts this is accommodated by simulating time of dropout for each hypothetical subject and applying the imputation method (e.g., last-observation carry forward) to the simulated data. This of course requires the development of a dropout (trial execution) model. The data analysis is then performed for each simulated trial for each design and the decision criteria are applied to make a decision (e.g., go or no go). This decision can be compared against the correct decision under the models and true values of the parameter used to simulate each trial. The probability of a correct decision can then be computed as one of the measures of trial performance. This will be discussed further in Section 2.6 (Trial Performance Metrics).

When the primary goal of the study is to understand the dose-response relationship to select doses for a phase 3 trial, a design may require a regression-based data analysis approach where many dose groups may be studied. With the new paradigm, the rationale for the specification of Δ as well as the variability and uncertainty in the prediction are made based on all the available data, so that a more realistic assessment about the probability of success (positive trial regardless of whether a go decision is the correct decision) can be determined for each potential design under consideration.

2.4 Data-Analytic Models

The formal, prospectively defined statistical analysis methods, with their underlying statistical models are referred to here as *data-analytic models*. For confirming questions (e.g. "Is the effect size for this treatment big enough?"), the best choice may be a simple pairwise ANCOVA comparison, which requires minimal assumptions (e.g., that randomization occurred, observations were blinded, and any drop-out is non-informative), particularly when the result will serve as the basis for regulatory decision-making. In contrast, learning questions, such as "What is the shape of the dose-response?" may be best addressed using a regression-based approach. The complexity needed for the regression model will depend on a more detailed understanding of the question. For example, an endpoint (fixed time point) analysis to estimate the dose-response relationship may only require a simple linear or nonlinear regression model, while, a longitudinal analysis, or an exposure-response model, may require a more complex linear or nonlinear mixed effects analysis, with stronger assumptions about the form of the dose and time relationships. The trial objective and the choice of primary analytic model will also influence the design, particularly the number and placement of dose groups and timing of measurements.

2.5 Quantitative Decision Criteria

Clinical trials are designed to ask questions about the effects of the drug in the population and conditions under study. Drug development decisions (e.g., terminate vs. continue development) will be based on the answers obtained. The objective during the trial design phase is to craft a design and associated analysis plan and evaluation criteria that will maximize the chances of making the correct development decision, averaged, in some sense, over the range of potential drug candidates and conditions that might be encountered, taking into account the potential risks and rewards of each potential decision pathway. This is a daunting task, one which cannot likely (with present tools) be fully completed on a routine basis for every clinical trial conducted. However, we believe that current practices can be greatly improved taking a more quantitative approach during the design phase.

It is imperative that key trial questions and the associated decision criteria be explicitly and quantitatively defined, so that the trial can be 'optimized' to provide the necessary information. Study designs based on arbitrary criteria where the actual decision criteria are left vague run the risk that the selected design will be inefficient and perhaps uninformative regarding the actual decisions to be made. The success of a model-based drug development approach hinges on the project team's ability to agree upon the questions that each trial must address and a pre-specified set of *quantitative decision criteria*. Getting such agreement can be one of the most difficult challenges in setting a development strategy, in part because project team members might not be comfortable with specifying quantitative criteria, especially when the decision is dependent on multiple factors. The quantitative scientists (e.g., statisticians, pharmacometricians, etc.) may need to take the lead and propose quantitative criteria that can be discussed and debated openly among the project team members to achieve consensus. Such discussions also reveal what the various team members believe are important factors affecting their decision, which can lead to formulating alternative development strategies that otherwise might not have been considered. Further, these discussions and the eventual criteria help us identify which models must be developed to inform the study design and whether or not the models, study design, and analysis method are suitable for the intended purposes.

It is often helpful to work backwards, perhaps iteratively, starting with the ultimate product and regulatory requirements, to define the quantitative data-driven decision criteria to be used at various stages in the clinical development plan. The following checklist of general planning questions is adapted from Sheiner (1997), which can be a useful starting point for this process.

- (1) What do we need to know?
- (2) How sure do we need to be?
- (3) What assumptions are we willing to make (including what we already 'know')?
- (4) When do we need to know?

Clearly, clinical development will be inefficient and potentially misguided unless our trials appropriately measure what is needed, are adequately sized (neither too big nor too small), and are based on reasonable assumptions for the purpose of the trial. Moreover, asking a good question at the wrong time can also lead to inefficient clinical development, because either key information is not available when needed for a decision, or resources could be wasted if the information is only needed if the compound proceeds to late-stage development.

To more formally quantify what we need to know it is helpful to define terminology regarding the treatment effect (Δ). Here Δ denotes the true (unknown) treatment effect of the candidate drug (under a specified dose regimen, in a particular population) relative to some reference treatment (typically placebo). It is convenient to discuss this true effect size as a simple difference between population means, but in practice it might be a regression model parameter, perhaps from a mixed effects model. The trial objective is to make inference regarding the size of Δ .

We can define some useful categories or classifications of Δ . Two treatments are considered *numerically identical* if $\Delta=0$, whereas the candidate drug response is *numerically inferior or numerically superior* when $\Delta<0$ or $\Delta>0$, respectively (assuming for discussion that larger values are preferred). A clinically relevant difference can be defined by specifying upper and lower clinical similarity limits (i.e., LCSL, UCSL), such that two treatments are considered *clinically similar* when $\text{LCSL} \leq \Delta \leq \text{UCSL}$. Five distinct regions are formed by comparison of Δ to the numerical equality and clinical similarity limits as illustrated in Figure 3.

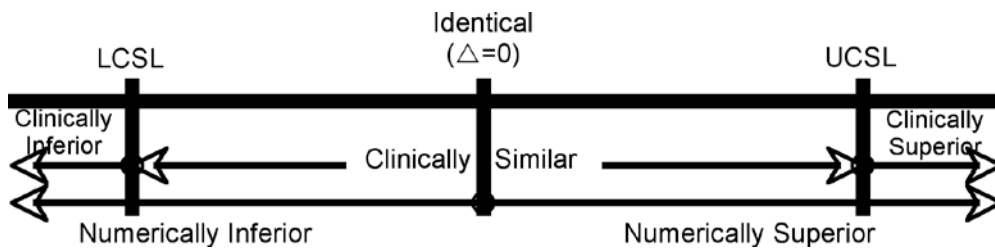


Figure 3. Possible Relationships of the True Response Between Candidate and Reference Treatments

Within the model-based drug development paradigm, planning trials around a statement about truth regarding Δ is accomplished by predicting Δ for a specified set of model assumptions. These model assumptions form the basis for a data-generation model to simulate trial data and employ CTS techniques to evaluate potential

designs and analysis methods. In so doing, both parameter as well as model uncertainty needs to be taken into consideration. For example, at an early stage in development multiple models may adequately describe the existing data but could lead to different predictions when extrapolating to different doses, regimens, or duration of study. Thus, we may have different sets of potential truths about Δ based on these different model assumptions. Therefore, for planning purposes it is important to evaluate designs under different sets of model assumptions to assess the sensitivity of the design to these different assumptions and to choose a design that is fairly robust to this model uncertainty.

To address the second planning question, “how sure do we need to be?” requires managing the risk of making the wrong development decision, e.g., passing a poor candidate drug (false positive) or terminating a clinically superior candidate drug (false negative). A way to manage these two risks is to employ statistical inference based on comparing the appropriate confidence intervals to the numerical equality and/or clinical similarity limits illustrated in Figure 3.

Consider the following example, in which we’ll assume the product requirement is that $\Delta > \text{UCSL}$, so that by the time of regulatory filing we will want strong evidence (e.g., $p < 0.025$) that $\Delta > \text{UCSL}$. It is impractical to expect this level of confidence during early clinical development, particularly if we want to have a high chance of bringing a candidate with Δ just above UCSL through to regulatory filing. A “dual criterion” decision rule can help us manage the false positive and false negative risks. To limit our risks of taking a clinically inferior drug into phase 3, we may wish to have a high level of confidence (e.g., 80%) that $\Delta > \text{LCSL}$. Further, we may want to limit the risk (e.g., to 10%) that we incorrectly terminate a clinically superior candidate (i.e., $\Delta > \text{UCSL}$); alternatively stated, an effect size $> \text{UCSL}$ should at least be ‘plausible’ (probability $> 10\%$) before we proceed to phase 3. An example decision rule that meets these requirements is as follows:

Go decision if: $\text{PCT}_{20} > \text{LCSL}$ and $\text{PCT}_{90} > \text{UCSL}$

Pause if: $\text{PCT}_{20} \leq \text{LCSL}$ and $\text{PCT}_{90} > \text{UCSL}$

Stop if: $\text{PCT}_{90} \leq \text{UCSL}$

where PCT_{γ} denotes the γ -th percentile of the distribution of the estimate of Δ ($\hat{\Delta}$), which might be from a single trial or a combined analysis of multiple trials, and either from a pairwise comparison or perhaps a regression model. We use this ‘PCT’ notation to avoid confusion between one-sided and two-sided lower and upper confidence limits, and ignore, for our pragmatic purposes, whether these percentiles are derived from a proper Bayesian posterior distribution or a fiducial interpretation of a conventional frequentist confidence interval. The focus of the criterion is whether there is evidence that the true effect is $<$ or $>$ UCSL (‘stop’ vs ‘go’), with a ‘gray zone’ result of ‘pause’ when the data are not sufficiently precise to confidently make a decision.

Figure 4 illustrates some possible outcomes of this decision rule. Each line represents a different study result, with the endpoints denoting the location of the 20th and 90th percentiles relative to the reference values LCSL and UCSL. A ‘stop’ recommendation is made in the 3 instances where the upper percentile falls below the UCSL, while a ‘go’ recommendation is made in the 2 instances where the lower and upper percentiles fall above the LCSL and UCSL, respectively. The ‘pause’ recommendation is made when the data are too imprecise for a confident decision (e.g.,

perhaps our pre-trial estimate of the variance was too small), with the idea that either more data will be collected before the stop/go decision is made or perhaps that additional endpoints might be evaluated to aid in making the stop/go decision.

It is not expected that such a decision rule would be slavishly followed, in part because gaining prior agreement regarding the precise similarity limits (LCSL, UCSL), and percentile levels will not always be possible. Furthermore, the recommendations from such a decision rule would need to be evaluated in the larger context of all the available information, clinical and otherwise, as part of the decision-making process. As we progress forward and complete further studies the criteria must gradually tighten towards the actual registration/target profile criteria.

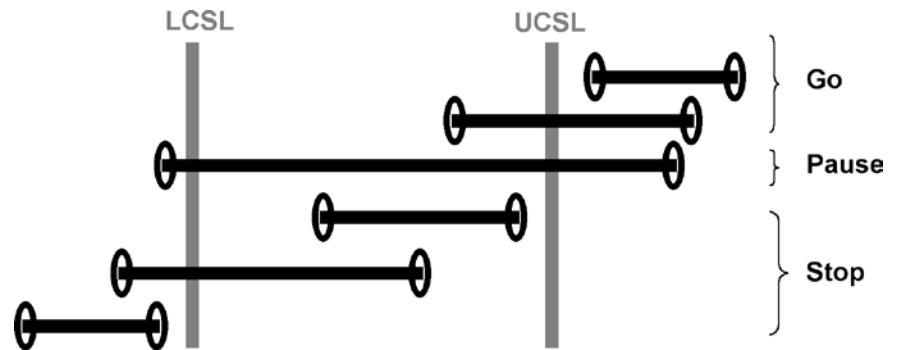


Figure 4. Example of a Dual Decision Criterion Rule

2.6 Trial Performance Metrics

In the simplest case of designing a two-group trial to test the null hypothesis $H_0: \Delta \leq 0$ vs $H_a: \Delta > 0$, the basic design issue is the size of each group (as well as, of course, the many context decisions such as patient population, length of treatment before the dependent variable is measured, etc). The trial performance metrics are also quite simple: We fix the desired type 1 error risk to some α , the probability of an incorrect decision when the true $\Delta=0$, and the type 2 error risk to some β , the probability of an incorrect decision when the true Δ is some target value, say Δ_0 . If we are willing to assume that we know the distribution and variance of the dependent variable, the required sample size can be easily calculated. This calculation is perhaps the simplest form of clinical trial simulation (CTS).

In clinical development, particularly in Phases 1 and 2, the questions which must be addressed in clinical trials and the framework for making decisions are much more complex than simply $H_0: \Delta \leq 0$ vs $H_a: \Delta > 0$. The best choices for trial operating characteristics – the best choices for the probability of a correct decision under various true states-of-nature – are usually not clear-cut, since factors such as corporate risk tolerance, the overall risk and rewards related to correct and incorrect decisions, and the opportunity costs related to other drug candidates in the development portfolio all must be considered. Further, assumptions such as $X \sim N(\mu, \sigma^2)$ might be overly simplistic when trying to evaluate trial performance.

CTS provides a useful basis for evaluating trial performance for a proposed study design. The following discussion illustrates how the six components of model-based drug development discussed in this article can be integrated within a CTS framework.

1. **Data models:** The data-generation models for CTS are derived from the relevant PK, PK/PD, disease, and trial executions models, which summarize the current knowledge base, and other assumptions appropriate for the context being considered. An important choice is whether the designs are to be evaluated under more-or-less arbitrary assumptions about the true Δ (e.g., $\Delta=0$ or $\Delta=UCSL$) or under the best current estimate of the true effects. The former yields something akin to a conventional power calculation, while the latter seeks to predict the actual trial outcome taking into account uncertainty in the prediction of the true Δ . In many cases, alternative models may fit the current data equally well and hence, different sets of assumptions about truth can be explored. These models provide ‘the truth’ against which the design performance can be assessed. CTS techniques can be used to evaluate the sensitivity to different assumptions so that a proposed trial design has acceptable operating characteristics across a range of plausible data models and assumptions.
2. **Designs:** The data generation model(s) need to be adapted to each design of interest. If this involves alternative dose groups or observation times, the models must be defined to support generating data under the variety of conditions.

3. **Analytic models:** One or more analysis methods might be considered and applied to each simulated trial, e.g., ANCOVA or a regression-based model, with appropriate adaptation to the particular design. Methods differ in the assumptions required and the information obtained from the data, potentially leading to different decisions for individual (simulated) trials and thus overall operating characteristics (e.g., the accuracy of estimates or the probability of a correct decision). The analyses must supply the quantities required by the proposed decision criteria.
4. **Decision criteria and drug development context (meta-analytic models):** The proposed decision criteria must be specified, possibly with alternative versions to be compared. So that the CTS recommendations are most relevant to the actual trial, these criteria should mimic the intended logic as closely as possible. The meta-analytic information, other clinical/commercial information, as well as corporate priorities and risk tolerance play a role in defining, for instance, in the context of the Section 2.5 example, the test statistics to be used (PCT_{20} , PCT_{90}), the reference levels (LCSL, UCSL), and the decision options to be considered (just Go vs No Go or perhaps additional 'gray zone' actions). Of course, every analysis does not feed directly into a major Go/No Go decision, so that a variety of criteria might be relevant at this step, for instance, achieving adequate precision in the estimated dose needed to provide a specified target effect size.
5. **Trial performance metrics:** The decision criteria are applied to the analysis results for each simulated trial (for each data-analytic model, design, and data-generation model combination), leading to the metrics of interest, whether this be a Go/No Go decision recommendation, an estimated target dose, or some other indicator of trial success or failure. Separately, based on the particular data-generation model, we will know what the correct answer is, so that each trial's answer can be scored (e.g., correct or not for a Go/No Go recommendation or sufficiently small error in a dose estimation). The entire process is then repeated a sufficient number of times so that the average metrics, over the replicate trials for each set of conditions, have sufficient precision to allow meaningful comparison of the design and analysis alternatives.

The design-evaluation process concludes with selecting a design, analysis model, and decision criterion that optimizes, in some sense, the chances that a trial will reach a correct decision or conclusion. This will inevitably require judgment to appropriately weight the performance under the various assumptions about the truth (i.e., alternative data models in Step 1) – typically we will want our selected design to be robust, in the sense of not performing too badly across the range of plausible realities, while being the best or near-best for the most likely conditions. Fundamentally, this is not much different from making a choice in the Type 1 vs. Type 2 error tradeoff, but the tradeoffs to be considered can be complex and have multiple dimensions when we are dealing with realistic decision frameworks, possibly involving both safety and efficacy decision criteria.

3 Examples

We now consider three real examples where modeling and simulation have played a prominent role in a Pfizer clinical development program. Without discussing the specific details of the models and analyses, we focus on how various model-based drug development concepts were useful for executing these programs.

The first example is SC-75416, a selective COX-2 inhibitor, which was in early clinical development for the treatment of acute and chronic pain. The second is CI-1017, a candidate for treatment of Alzheimer's disease (AD). The third example is referred to here as Drug X. Table 2 summarizes how the six components of the model-based drug development paradigm were applied in each example.

3.1 SC-75416

SC-75416 is a selective COX-2 (cyclooxygenase-2) inhibitor, being investigated for acute and chronic pain indications. Initially, the clinical development strategy was to pursue a chronic pain indication in osteoarthritis (OA) patients. A dose-ranging dental pain study was conducted to determine an acute pain dose that would give comparable pain relief to 50 mg rofecoxib (Vioxx™). In addition, an exposure-response meta-analysis was performed

Table 2: Examples of Model-Based Drug Development

Component	1. SC-75416	2. CI-1017	3. Drug X
PK/PD and Disease Model Development	Developed PK/PD model for pain relief (PR) scores in patients with acute pain following oral surgery, using potency scaling to pool data from other compounds in the same class and predictions for other formulations based on observed PK profiles.	Combined ADASCog dose-response model for tacrine, healthy volunteer PK for CI-1017, preclinical information for CI-1017, and plausible alternative concentration-response models (monotonic: linear, Emax, sigmoid-Emax; non-monotonic: inverted U-shaped).	Developed three alternative longitudinal dose-response models (K_{out} , Direct, K_{eo}), based on a phase 2A POC study, allowed extrapolation to lower doses and longer duration study.
Meta-Analytic Model of Competitor Data	Exposure-response meta-analysis was performed across a number of marketed NSAIDs and Cox-2 inhibitors to estimate relative potencies between WOMAC-pain in OA and dental pain TOTPAR6 responses. Meta-analytic model was used to scale between acute (dental pain) and chronic (OA) pain doses.	Competitor data was considered in setting effect size and time-to-onset product requirements	Competitor data was analyzed and considered in setting response target values.
Design Considerations and Trial Execution Model	Considered parallel group designs with various doses and sample sizes, and active comparator. Utilized a non-informative dropout model that depends on the most recent PR score.	Considered parallel group, Latin-square, and incomplete block designs.	Considered parallel group designs with alternative size and number of dose levels.
Data Analytic Model	ANOVA of TOTPAR6 responses.	ANOVA or ANCOVA of ADASCog responses (design-dependent).	Three parameter Emax dose-response logistic regression model for the binary endpoint.
Decision Criteria	Pairwise comparisons of $\Delta_{TOTPAR6}$ relative to active. Go decision if $PCT_{2.5} > 0$ ($LCL_{95} > 0$).	Pairwise comparisons of $\Delta_{ADASCog}$ relative to placebo. Go decision if $PCT_{2.5} > 0$ ($LCL_{95} > 0$). Dose-response shape inferred from observed data patterns.	Pairwise comparisons of responder rates (Y) relative to placebo. Go decision if $PCT_{10} > 20\%$ and $PCT_{90} > 30\%$ ($LCL_{80} \geq 20\%$ and $UCL_{80} > 30\%$)
Trial Performance Metrics	Probability of Go decision for various designs (doses, sample size).	Power for positive trial and probability of correctly detecting shape of the dose-response relationship.	Probability of a correct decision: Go when true $Y > 30\%$ No Go when true $Y \leq 30\%$

for several marketed NSAIDs and COX-2 inhibitors to estimate the relative potency between WOMAC-pain (OA) and TOTPAR6 (dental pain) responses. From this meta-analysis, we estimated that chronic doses to treat OA pain are generally about 2- to 4-fold lower than acute pain doses. Our plan was to determine the acute pain dose that achieves pain relief comparable to rofecoxib and use the meta-analytic model to estimate comparable doses for OA. This allowed us to leverage the easier to obtain and less costly dose-response information from the dental pain setting. Scaling this acute pain dose-response to the OA setting based on the meta-analytic model, allowed us to rationally select a narrower range of OA dose(s) than otherwise possible.

The parallel group dental pain study evaluated single oral doses (placebo, 3, 10 and 60 mg of SC-75416, and 50 mg rofecoxib), with 50 patients per group. Pain relief (PR) scores on a 5-point Likert scale (PR=0: no pain relief; PR=4: complete pain relief) and SC-75416 plasma concentrations were measured at several times over 24 hours.

Patients could drop out of the study and take rescue medication after one hour. Missing PR scores after dropout were imputed by last observation carried forward (LOCF). The scheduled SC-75416 plasma concentrations were collected regardless of rescue medication use. From preclinical potency estimates and PK models from healthy volunteer studies, we expected that the highest SC-75416 dose (60 mg) would achieve pain relief comparable to 50 mg rofecoxib. The results demonstrated statistically significant analgesic efficacy relative to placebo for the 10 and 60 mg doses, but no dose achieved pain relief comparable to 50 mg rofecoxib. Evaluation of the PK data revealed that the dental study patients (capsule formulation) had slower and more erratic absorption, particularly at the critical early time points (up to 2 hours), compared to the healthy volunteers who had taken an oral solution. We surmised that the poor pain relief response was due to lower than planned plasma concentrations achieved by the capsule formulation. This raised the question: How would a 60 mg oral solution dose have performed in the dental pain study? We developed a PK/PD model to address this question, as a way to salvage useful information from the otherwise flawed study.

The PK/PD model related plasma concentrations of SC-75416 and rofecoxib to the pain relief scores, using a nonlinear mixed effects logistic-normal model. A dropout model related the dropout hazard to the last available PR score; this requires the assumption that dropouts are missing at random (see Sheiner, Beal and Dunne, 1997). These models predicted that a capsule with a PK profile similar to the 60 mg oral solution would have had pain relief comparable to 50 mg rofecoxib. Moreover, extrapolated predictions suggested that higher doses could achieve clinically superior pain relief relative to 50 mg rofecoxib and other marketed NSAID and COX-2 inhibitor products. This intriguing hypothesis encouraged the team to consider pursuing an acute pain development strategy, with the high dose providing efficacy differentiation from the other products.

We proposed a second dental pain study, using an oral solution formulation, to evaluate doses up to 360 mg SC-75416 compared to rofecoxib; however, due to rofecoxib's withdrawal from the market, we instead used 400 mg ibuprofen as the active comparator. This required additional PK/PD modeling to pool dental pain data from another Pfizer project, valdecoxib (Bextra®), where 400 mg ibuprofen had been used as an active control. The updated PK/PD and dropout models provided the data-generation models for clinical trial simulations that evaluated alternative designs (doses and sample sizes) for an oral solution dental pain study which would test our hypothesis that a high dose of SC-75416 could achieve numerically superior pain relief to 400 mg ibuprofen. TOTPAR₆, a time-weighted sum of the LOCF PR scores over 6 hours, was proposed as the primary measure of pain relief. The analytic model for TOTPAR₆ used ANOVA for pairwise comparisons between the various doses of SC-75416 and 400 mg ibuprofen. The decision criterion to continue pursuing this high dose acute pain strategy was: 'Continue' if $PCT_{2.5} > 0$ for $\Delta_{TOTPAR_6} = TOTPAR_{6(SC-75416)} - TOTPAR_{6(ibuprofen)}$, where $PCT_{2.5}$ denotes the lower limit of the 95% CI for Δ_{TOTPAR_6} . We simulated 1000 trials for each design considered, with the trial performance metric being the percentage of 'successful' trials, i.e., with $PCT_{2.5} > 0$, evaluated on a per-dose-level basis. ANOVA was used as the data-analytic model since we considered this as a confirmatory test of the hypotheses generated from the PK/PD modeling.

We considered a $\Delta_{TOTPAR_6} = 3.0$ as a clinically relevant improvement in pain relief relative to 400 mg ibuprofen. Our models predicted that 360 mg SC-75416 would provide this level of efficacy (predicted $\Delta_{TOTPAR_6} = 3.2$). To have at least 80% probability to obtain a statistically significant comparison if the true difference were 3.0, 100 patients per treatment arm would be needed. We also were interested in evaluating the dose-response, particularly over the range from 60 to 360 mg. On the basis of the clinical trial simulations, we planned and executed a 5-arm parallel group dental pain study evaluating placebo, 60, 180 and 360 mg SC-75416 oral solution, and 400 mg ibuprofen. The primary comparison of interest was 360 mg SC-75416 vs 400 mg ibuprofen, each with 100 patients, while the remaining groups had 50 patients.

The study results confirmed our model-generated hypothesis, with a statistically significant ($PCT_{2.5} > 0$) difference of $\Delta_{TOTPAR_6} = 3.3$ for 360 mg SC-75416 oral solution compared to 400 mg ibuprofen. Moreover, the observed means were well within the predictive distribution (obtained from the clinical trial simulations) for each active treatment, substantiating the predictive performance of the PK/PD and dropout models.

This model-based development strategy allowed progress to be made in understanding the exposure-response relationship for this compound without having to wait for a new solid dosage form to be developed, a time savings of approximately 9 months. Reformulation work was done in parallel and the 'validated' PK/PD model provided

the team with the confidence to evaluate the performance of potential new formulations on the basis of PK data alone. Thus, the team was prepared to pursue the use of a new formulation in a different acute pain model without having to repeat the dental pain study with the new formulation, a potential cost-savings of \$1.5 M and a 7 month time savings. However, the greatest impact this modeling and simulation effort had on the program was in providing the rationale for pursuing the high dose strategy and designing a study to test the efficacy differentiation hypothesis that might not have otherwise been considered.

3.2 CI-1017

CI-1017 is an M₁-muscarinic acid agonist that was under investigation for the treatment of Alzheimer's disease (AD). The Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADASCog) is a primary endpoint in AD trials typically determined at 4-8 week intervals over a 12-30 week trial duration. Substantial clinical trial data was available for tacrine, an approved treatment for AD symptoms. The tacrine data provided the minimum efficacy profile for a drug that would be worthy for further development, namely one that would provide a $\Delta_{\text{ADASCog}} \geq 3$ points relative to placebo, achieved in approximately 10-12 weeks.

The planning of the first efficacy trial was complicated by the suggestion from preclinical models that the dose-response relationship might be non-monotonic ("U-shaped"). So that besides the basic proof-of-concept efficacy question for this trial, it was also important to learn about the shape of the dose-response relationship, both to allow choice of doses for subsequent phase 2B/3 trials and to assess the viability of the compound: If it were active but with a non-monotonic dose relationship, we would need to search for the dose providing the best effect. Moreover, since individuals would likely differ in their maxima, some form of individual dose-titration to response would likely be required, probably rendering the drug practically and commercially nonviable. Thus, we needed this trial to focus on assessing the shape of the dose-response relationship, as well as on the primary objective of whether any dose of CI-1017 improves ADASCog.

The originally proposed design had 6 parallel groups, placebo and 5 dose levels of CI-1017, with efficacy assessed after 12 weeks, and approximately 60-80 patients per group – a large, expensive trial for a candidate at the POC stage. The project team explored other more feasible design options, one being the use of a 6 treatment, 6 period Latin Square, using 2 week treatment periods and no washout between periods. Initial power calculations suggested that a total of about 60 patients would be sufficient to answer our primary questions as well as the 360-480 patient trial originally proposed. However, many team members were rightfully skeptical, since this situation is one where cross-over designs are contraindicated: patients will not reach their full treatment effect in 2 weeks and there would most certainly be (perhaps substantial) carry-over of effect from one period to the next.

We conducted a CTS project to assess how various designs compared, to assure ourselves that the proposed Latin Square design and analysis would be able to meet our objectives. We compared the original parallel group design and several variations of Latin Square and Incomplete Block designs (using design-specific analysis methods) for their ability to detect a treatment effect and to characterize the dose-response relationship as either monotonic or U-shaped; unbiased estimation of the treatment effect size was not a primary objective. We used data generation models that incorporated the slowest-acceptable time pattern for onset and offset of treatment effects (thus inducing carry-over in the cross-over trials). Several alternative dose response models were employed (monotonic: linear, Emax, sigmoid Emax; U-shaped), each reaching the minimally required effect size in the proposed trial dose range (as well as a no-effect model to confirm type 1 error). With these basic features set, we wanted the simulated data to be as realistic as possible, considering temporal effects and both systematic (demographic) and random individual variation in model parameters. The starting point for the disease progression and dose-response models was the work by Holford and Peace (1992a, 1992b, 1994). This was combined with the phase 1 and preclinical PK/PD data for CI-1017.

The two key trial performance metrics were the percentage of trials that detected a drug effect ($\text{PCT}_{2.5} > 0$ – essentially the conventional statistical power), and the percentage of trials that correctly identified the shape of the dose-response (monotonic vs U-shaped, based on the pattern of least squares means). For more details, see Miller et al. (2004).

The CTS results suggested that a 4-period crossover design with 4-week periods was the most robust and performed well (>80% power) for all three monotonic dose-response relationships. The power was lower for the

U-shape pattern but higher compared to the other designs. The parallel group design, which was constrained to have the same total number of subjects ($N=60$), did not perform well, confirming the original power calculations that hundreds of patients would be needed to have comparable power to the Latin Square design. The results for the second objective, detecting the shape of the dose-response relationship, were qualitatively similar.

Based on these CTS results, the 4-treatment 4-period crossover design, with 4-week periods and no washout, was recommended and ultimately conducted. The trial results did not meet the predefined 'go criterion' and the CI-1017 program was terminated.

The clinical trial simulations were helpful in facilitating discussions with the project team concerning objectives, what was known and/or could be assumed about AD and CI-1017, and the uncertainty in our knowledge. Moreover, the CTS results helped the team select and receive support for the potentially controversial crossover design, by providing tangible 'evidence' of its feasibility and operating characteristics. The study used a fraction of the traditionally required number of patients, saving more than \$2 M relative to a parallel group design with similar power. Moreover, a time savings of approximately 8 months was realized because of the reduced enrollment burden. Finally, because the design was considered robust to various possible dose-response relationships, the project was cleanly terminated without the need for endless rework of the data to look for an elusive signal as is often the case with negative studies.

3.3 Drug X

The third example involves a drug in clinical development which will be referred to as Drug X. A phase 2A proof of concept (POC) parallel group study was conducted comparing 5, 15 and 30 mg BID doses to placebo over 6 weeks, with approximately 60 patients per group. The primary efficacy endpoint, Y , is a binary response ($Y=1$ denotes responder, $Y=0$ denotes nonresponder). Statistically significant ($p<0.05$) improvements in the responder rates were achieved with observed mean $Y > 40\%$ relative to placebo for all three dose groups. However, adverse events were also observed, particularly at the highest dose (30 mg BID) that could limit its usefulness in longer duration therapy. The project team was interested in exploring lower doses (15 mg or less) in a longer duration (12-26 weeks) phase 2B study, aimed at selecting a dose(s) for phase 3.

Longitudinal dose-response models were developed for both efficacy and safety measures, for use in clinical trial simulations to assess dose selection and sample sizes for the proposed phase 2B study. A key objective of the modeling and simulation effort was to determine whether the proposed study would be capable of identifying a 'therapeutic window' of safe and effective doses, based on different model assumptions of what the true underlying response profiles are. To illustrate some of the model-based drug development concepts, we will only focus on the efficacy measure.

Three different longitudinal logistic-normal mixed effects dose-response models were developed; each adequately fit the POC study data. Specifically, the binary responses were fit using a nonlinear logistic regression model in which additive placebo-time and drug fixed effects, and inter-individual normal random effects were postulated in the logit domain. We refer to these three dose-response models as the 'direct', ' k_{eo} ' and ' k_{out} ' models. The direct model assumes that the drug effect follows an Emax model using dose as the measure of exposure; the drug effect is instantaneous and the time-course of response is attributed solely to the placebo effect. The k_{eo} model assumes that there is a temporal delay in the drug effect due to the distribution of the drug to an effect site and hence the full effect of the dose is dependent on the equilibration rate constant (k_{eo}); both the placebo and drug effects contribute to the time-course of the response. The k_{out} model assumes that temporal delay in the drug effect is due to an indirect response mechanism, and like the k_{eo} model, both the placebo and drug effects contribute to the time-course of the response. The indirect response model is the most biologically plausible based on the current understanding of the mechanism of action of the drug. Nevertheless, the POC study data do not provide sufficient information to rule out any of these models and they all give similar predictions over the dose range (5-30 mg) and duration (≤ 6 weeks) studied. However, the models yield different predictions when extrapolating to lower doses (< 5 mg) and later time points (> 6 weeks). To ensure that our chosen design would have good operating characteristics over a range of true relationships; all three models were used as data-generation models in the clinical trial simulations.

The ultimate efficacy target for this drug is that the true responder rate be at least 30% relative to placebo. To balance false positive and false negative risks from an efficacy perspective, an 80% confidence interval (PCT_{10} ,

PCT₉₀) criteria on the responder rate (Y) was proposed for the phase 2B trial, such that a go-to-phase 3 recommendation would be made if a safe dose could be found with the efficacy response being PCT₁₀ > 20% and PCT₉₀ > 30%; this assessment would be based on regression models so that a recommended phase 3 dose need not necessarily have been included in the phase 2B trial. We considered a range of phase 2B designs of longer duration (>6 weeks) exploring lower dose levels (≤ 15 mg) of Drug X and sample sizes. For each data-generation model and design combination, 500 clinical trials were simulated. For each simulated trial, a nonlinear logistic regression model using a three-parameter Emax model (including intercept term) was fit to the binary response data at the last study visit. This data-analytic model requires fewer model assumptions than the data-generation models and appears to have sufficient flexibility to describe the dose-response at a fixed time point regardless of the true underlying longitudinal dose-response model (among the three data-generation models used). Based on the parameter and covariance estimates for the data-analytic model, parametric bootstraps were performed to calculate the 80% confidence intervals (PCT₁₀, PCT₉₀) about the responder rate prediction for each dose group in the simulated trial. These confidence limits were compared to the decision criteria.

The trial performance metrics were summarized over the 500 replicate trials as shown in Table 3, for each combination of design (doses and sample sizes) and data-generation model (assumptions about the true dose-response relationship). A key feature of this simulation set-up is that we incorporated the uncertainty about the data model parameters (covariance matrix from fitting the model to the original data) into the simulated data. Thus, at some particular dose, it might have the desired efficacy in one simulated trial ($Y > 30\%$) while being inadequate ($Y < 30\%$) in another. Therefore, the desired 'correct' decision at each dose must be determined and accounted for in each trial. The summary of the 500 trials in Table 3 is represented by 2x2 contingency table classifications of 'truth' (or desired decision) versus data-driven (observed) decision.

Table 3: Proposed Phase 2b Trial Performance Metrics for Drug X

Truth (Desired Decision)	Data-Driven Decision		Total
	Go (Success) PCT ₁₀ >20% \cap PCT ₉₀ >30%	Stop (Fail) PCT ₉₀ \leq 30%	
Y >30% (Go)	Prob(Go \cap Y >30%)	Prob(Stop \cap Y >30%)	Prob(Y >30%)
Y \leq 30% (Stop)	Prob(Go \cap Y \leq 30%)	Prob(Stop \cap Y \leq 30%)	Prob(Y \leq 30%)
Total	Prob(Go)	Prob(Stop)	1.0 (500 trials)

These summary data can be aggregated in two ways: first, we might consider the conventional probability of trial 'success' i.e., a go decision. Second, we should be most interested in the probability of a 'correct' decision, i.e., of either a stop if $Y \leq 30\%$ or a go if $Y > 30\%$, which is the sum of the upper left and lower right cell in Table 3. Note that this definition of 'correct' makes a very specific and precise distinction between the desired decisions for candidate doses with effects on either side of the reference value of 30%, so this value must be chosen with care.

In addition, the simulation tracked the accuracy of the dose estimation for the dose predicted to achieve $Y = 30\%$. This was summarized by the percentage of trials in which the estimate was within ± 2 and ± 3 mg of the true dose (based on predictions from the relevant data-generation model).

The full value of this modeling and simulation effort to the Drug X program is not clear yet, as the phase 2B study protocol is presently still under final development. However, the process helped facilitate team discussions, leading to a set of quantitative decision criteria, which will guide the eventual data evaluation. Moreover, the extensive data analysis during the modeling (and simulation) phases helped to determine our best current estimate of the lowest viable dose from an efficacy perspective. This information aided the decision of whether the currently available dose range was sufficient or whether further formulation development for lower doses would be needed (thereby delaying the trial start). All three dose-response models suggested that doses lower than the smallest dosage currently available would be nonviable for efficacy so that proceeding with the currently available supplies was appropriate, and of course, much more expedient.

4 Discussion

We have outlined components of a model-based drug development approach and provided selected examples illustrating their application and positive impact upon these development programs. From our observation, this approach is not routinely applied in drug development. This in part is due to the gradual and relatively recent diffusion, into the pharmaceutical development world, of the underlying principles from other industries where modeling and simulation are frequently used. Further, because pharmaceutical work, and the underlying pharmacology, is so complicated and varied, there are many challenges to 'industrialization' of the model-based drug development paradigm that must be addressed if we are to maximize its potential. Our concluding discussion highlights some of these challenges, particularly for the three major scientific disciplines within a clinical development organization, clinicians, clinical pharmacologists, and statisticians.

For *clinicians*, the main challenge is becoming comfortable with explicitly and quantitatively defined decision rules. In this regard, we reemphasize that any set of criteria would not be slavishly followed to the exclusion of other information bearing on the decision to be made. For this reason, the term 'decision recommendation' might be better, implying that the actual decision will be made within the larger context of all available information. That being said, if the criteria are vague the development strategy may be inefficient, and the data generated may be less informative than it could be otherwise. In our experience, the quantitative scientists on project teams often need to take the lead by proposing an initial version of the criteria that can be debated openly. With consensus achieved around the decision rules, the team can focus on any knowledge gaps that must be addressed.

One of the key challenges for *clinical pharmacologists* and *pharmacometricians* working in quantitative drug development is to be explicit and transparent about the assumptions and limitations of the PK/PD and disease models they develop. Vague descriptions of the assumptions and limitations can lead to misunderstanding and skepticism, and make it difficult to assess whether the models are truly fit for purpose. Open discussion provides opportunity for greater collaboration and eventually better buy-in for exploiting these models to inform strategy, trial designs, and decision-making.

To help build trust that a model is adequate for its intended purpose, it is helpful to calibrate it against data-derived (non-model-based) statistics of interest in each particular situation. In this process, we use the models to simulate data, apply the conventional statistical analysis to the simulated data, and then show the statistical quantities of interest (e.g., treatment effects, mean squared error from an ANOVA, etc.) are similar to those actually observed. To improve communication, pharmacometricians must learn to avoid excessive use of technical jargon when communicating to a broader audience.

For *statisticians* a key challenge is learning to embrace assumption-rich nonlinear models, particularly for internal decision-making especially in early clinical development – i.e., avoiding the 'mini-phase 3' mentality when designing phase 2 studies. Reliance on empirical, assumption-poor models to make decisions in early clinical development can be costly, both in the size of trials required and the accuracy of decisions made. Phase 2 learning questions may require different study objectives, designs, and analysis approaches than phase 3 confirming questions. Finally, statisticians should seek out opportunities to collaborate with clinical pharmacologists and pharmacometricians to develop new and innovative statistical methodology for model development and their strategic application to trial design and drug development decision-making.

The greatest challenges to the routine implementation of quantitative drug development may rest in the hands of *senior management*. Organizational structures and processes in the pharmaceutical industry have been optimized for a 'go fast at risk' strategy. Much emphasis has been placed on streamlining the mechanical processes for efficient conduct and reporting of individual trials. The processes and organizational structures may need to be revised to 'industrialize' model-based drug development. Senior management must understand that this is an 'investment in knowledge' strategy, so that sufficient time (between studies) and resources may be required to maximize the learning before making a decision or embarking on the next trial. Faster is not always better, as it is easy to make 'bad' decisions quickly. Hopefully offsetting any increased between-study time and increased staffing is that greater efficiency and success will be obtained by conducting the appropriate studies that address the right questions. It will take courage on senior management's part to trust the organization's scientists to fully embrace this new drug development paradigm.

Models essentially are the knowledge repository for the data that is generated in our clinical development programs. The strategic application and exploitation of the knowledge contained in these models can lead to more efficient and cost-effective drug development. Organizations that embrace model-based drug development as an organization-driven activity, and not just a value-added activity that is sometimes employed when it is 'perceived' as adding value, will have a distinct competitive advantage.

5 References

- Berry, D.A., Muller, P., Grieve, A.P., Smith, M.K., Parke, T., Blazek, R., Mitchard, N., and Krams, M. (2002). "Adaptive Bayesian designs for dose-ranging trials." In *Case Studies in Bayesian Statistics V* (Edited by Carlin, B., Carriquiry, A., Gatsonis, C., Gelman, A., Kass, R.E., Verdinelli, I., and West, M.). Springer Verlag, 99-181.
- Davidian, M. and Giltinan, D.M. (1995). *Nonlinear Models for Repeated Measurement Data*. Chapman and Hall.
- Dayneka, N.L., Garg, V., and Jusko, W.J. (1993). "Comparison of four basic models of indirect pharmacodynamic responses." *J. of Pharmacokinetics and Biopharmaceutics*, 21, 457-477.
- Food and Drug Administration (2004). *Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products*. www.fda.gov/oc/initiatives/criticalpath/whitepaper.html.
- Gibaldi, M., and Perrier, D. (1982). *Pharmacokinetics* (2nd ed.). Marcel Dekker.
- Holford, N.H.G., and Sheiner, L.B. (1981). "Understanding the dose-effect relationship: Clinical application of pharmacokinetic-pharmacodynamic models." *Clinical Pharmacokinetics*, 6, 429-453.
- Holford, N.H.G., and Peace, K.E. (1992a). "Methodologic aspects of a population pharmacodynamic model for cognitive effects in Alzheimer's patients treated with tacrine." *Proceedings of the National Academy of Sciences, USA*, 89, 11466-11470.
- Holford, N.H.G., and Peace, K.E. (1992b). "Results and validation of a population pharmacodynamic model for cognitive effects in Alzheimer's patients treated with tacrine." *Proceedings of the National Academy of Sciences, USA*, 89, 11471-11475.
- Holford, N.H.G., and Peace, K.E. (1994). "The effect of tacrine and lecithin in Alzheimer's disease: A population pharmacokinetic analysis of five clinical trials." *European J. of Clinical Pharmacology*, 47, 17-23.
- Kastrissios, H., and Girard, P. (2003). "Protocol deviations and execution models." In *Simulation for Designing Clinical Trials* (Edited by Kimko, H.C., and Duffull, S.B.). Marcel Dekker, 55-72.
- Mandema, J.W., Hermann, D., Wang, W., Sheiner, T., Milad, M., Bakker-Arkema, R., and Hartman, D. (2005). "Model-based development of gemcabene, a new lipid-altering agent." *The AAPS Journal*, 7, E513-E522 (<http://www.aapsj.org>).
- Miller, R., Hermann, D.J., Ewy, W., and Kowalski, K.G. (2004). "The strategic role and application of pharmacokinetic/pharmacodynamic modeling in drug development." In *Pharmacokinetics in Drug Development: Clinical Study Design and Analysis Vol. 1*, (Edited by Bonate, P.L., and Howard, D.R.). AAPS Press, 551-581.
- Sheiner, L.B., Stanski, D.R., Vozeh, S., Miller, R.D., and Ham, J. (1979). "Simultaneous modeling of pharmacokinetics and pharmacodynamics: Application to d-tubocurarine." *Clinical Pharmacology and Therapeutics*, 25, 358-371.
- Sheiner, L.B. (1997). "Learning versus confirming in clinical drug development." *Clinical Pharmacology and Therapeutics*, 61, 275-291.

Sheiner, L.B., Beal, S.L., and Dunne, A. (1997). "Analysis of nonrandomly censored ordered categorical longitudinal data from analgesic trials." *J. of the American Statistical Association*, 92, 1235-1244.

Smith, M.K., Jones, I., Morris, M.F., Grieve, A.P., and Tan, K. (2006). "Implementation of a Bayesian adaptive design in a proof of concept study." *Pharmaceutical Statistics*, 5, 39-50.

Vonesh, E.F., and Chinchilli, V.M. (1997). *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. Marcel Dekker. ■

Biopharmaceutical Section Members Elected ASA Fellows

For nearly a century, the American Statistical Association has recognized outstanding members of the statistics professions with the honorary title "Fellow of the American Statistical Association," or simply ASA Fellow. This is a great honor, limited each year to no more than one third of one percent of the ASA membership. From the ASA By-laws, "By the honorary title of Fellow the Association recognizes full members of established reputation who have made outstanding contributions in some aspect of statistical work."

At the 2007 Joint Statistical Meetings in Salt Lake City, seven members of the Biopharmaceutical Section will be given this honor. We are delighted to see them honored in this way. Their names are:

Scott S. Emerson, Professor of Biostatistics: University of Washington

Edward Lakatos, President: BiostatHaven, Inc.

Katherine L. Monti, Director: Rho, Inc.

Walter W. Offen, Senior Research Fellow: Lilly, Inc.

Ming T. Tan, Professor: University of Maryland

Russell D. Wolfinger, Dir. of Scientific Discovery and Genomics: SAS Institute, Inc.

Ji Zhang, Vice President Clinical Operations: Sanofi-Aventis, Inc.

Ever wondered how a statistician gets named ASA Fellow? In short, the process starts with nomination of an ASA member by any member of ASA, consisting of a nomination package submitted by March 1st to the ASA Committee on Fellows. The Committee makes its decision based upon positions held, publications, contributions to ASA, and other contributions to the profession. Letters of support are an important part of each nomination package.

Ever wondered whether one of your colleagues is an ASA Fellow? The ASA website maintains a list of every ASA fellow, dating from 1914. Do you know someone who is deserving of the honor, but is not yet on the list? You can find particulars on submitting a nomination on the ASA website (type ASA Fellow in the search box on the main page). Anyone, not just ASA Fellows, can submit a nomination.

As a service to members of the Biopharmaceutical section, our section has a Fellows Committee. We do not have any relationship to the ASA Fellows Committee, and we don't prepare nomination packages ourselves. Our contribution is simply to encourage nomination of worthy members of the section and to facilitate strong nomination packages. We will also write a letter of support for nominees who have made outstanding contributions to the Biopharmaceutical Section. For 2006 the committee consisted of Joe Shih, Keith Soper, and Bob Starbuck. In 2007 Joe Shih is leaving our committee and Greg Campbell is joining us. If you know someone who should be a Fellow, let us know! We need to identify candidates now to develop strong nomination packages. You don't have to commit to preparing a nomination package yourself.

If you are able to attend JSM this year, please attend the ceremony and offer your congratulations in person. Their accomplishments and recognition serve to enhance our profession as well as their own reputation. ■

Highlights of the Biopharmaceutical Section Executive Committee Meeting, Atlanta, Georgia, March 12, 2007

Brian Wiens

Following are highlights of the Biopharm Section Executive Committee Meeting in Atlanta, Georgia, on March 12, 2007.

Mani Lakshminarayanan gave the treasurer's report. The Section has a healthy balance, including a small profit from the 2006 FDA-Industry Workshop. A long discussion on ways of using Section money to help members was initiated. Any Section member is encouraged to contact any member of the Executive Committee with suggestions.

Dionne Price, FDA-Industry Workshop co-chair, reported that the annual Workshop will be held September 17-19 in Alexandria, Virginia. Topics and speakers are being determined. Registration should open in early summer. Matilde Sanchez is the other co-chair.

Phil Pichotta, *Biopharmaceutical Report* editor, reported that the winter issue should be ready in the next few days. He and the rest of the editorial team (Rick Caplan and Tom Dobbins) have ideas for feature articles for future issues. Any Section member with an idea for a feature article or for any other article is encouraged to contact a member of the editorial team.

Shuguang Huang reported on the balloting for the Best Contributed Paper Award. He will begin contacting winners from 2006 in the coming days. Volunteers are needed for JSM2007 to distribute and collect ballots at contributed paper sessions. Anyone interested can contact Shuguang.

The future of the Membership Committee was discussed. This committee has temporarily disbanded until new members can be identified. The Executive Committee agreed unanimously that the Membership Committee should be formed again, with a clear mandate and charter. The most important step is to identify a chair of the committee.

The new webinar series has been well-received. Rick Peterson of the ASA office has been extremely helpful in navigating the ASA bureaucracy and assessing the technology. The first webinar on Multiple Comparisons by Alex Dmitrienko is scheduled for March and is followed by a series of four webinars by Geert Verbeke and Geert Mohlenberghs in April – June. These webinars are meant to be a convenient and cost-effective training option, with a small subsidy from the Section. Departments are encouraged to purchase a single registration (pay only one fee for one connection) and project the webinar in a conference room to make the webinars even more cost-effective and better simulate a classroom. Feedback is encouraged.

Russ Helms joined for a discussion of how the Section and the Drug Information Association can work together. Russ is a co-chair of the statistics subgroup of DIA. Russ agreed to serve as an official liaison between the Section (of which he is a member) and the DIA statistics subgroup to aid in communication and cooperation.

Jim Colaianne reported on activities of the Corporate Sponsors committee. Due to a clerical error, the wrong address was included in the solicitation letters. Corrections are being developed and it is anticipated that due to the error the sponsorship money might be received later than in previous years.

Keith Soper reported on activities of the Fellows Committee. He, Bob Starbuck and Joe Shih identified several Section members to support for nomination to the title of Fellow of the ASA. Keith is looking for more members of the Fellows Committee. Anyone who is interested in nominating a Section member for Fellow should contact Keith.

Neal Thomas reported on activities of the Publications Officer. He has been successful in submitting articles for publication in *Amstat News* most months. The Executive Committee agreed that we should use the monthly column in *Amstat News* to publicize Section activities and communicate with members.

The Section is co-sponsoring the Graybill Conference in 2008. This is an annual conference in Fort Collins, Colorado. The 2008 Conference will focus on pharmaceutical statistics, and is scheduled for the second week of June. Naitee Ting, Fred Balch and Brian Wiens are co-chairing this meeting.

Amit Bhattacharyya and Kalyan Ghosh reported on progress for the 2007 and 2008 programs, respectively. Amit has the 2007 program nearly complete, including continuing education courses, roundtable luncheons, contributed papers and invited papers. Kalyan will be looking for proposals for invited paper sessions for JSM2008 very soon, since the final choices are determined soon after JSM2007 concludes, and will look for other activities in the near future. ■

Biopharmaceutical Section Program for JSM 2007

Amit Bhattacharyya

The 2007 Joint Statistical Meeting is coming up soon. Biopharmaceutical Section members are encouraged to start planning to attend this conference. This year, the program is hosted from July 29 to Aug 2 in Salt Lake City, Utah. This is one week earlier than usual. The meeting starts on Sunday (July 29) afternoon and finishes on Thursday afternoon (Aug 2). The continuing education classes start on Saturday (July 28).

The Biopharmaceutical section, being one of the largest sections of the ASA, is sponsoring and co-sponsoring a record number of invited, topic contributed, and contributed sessions at the conference. This year, there are 7 Invited sessions, 26 Topic Contributed session and 20 Contributed sessions primarily sponsored by the section. The topics encompass discovery and early to late phases of drug and device development. The section is sponsoring 13 roundtable luncheons (4 on Monday, 5 on Tuesday and 4 on Wednesday) on pharmaceutical and devices related topics. For the benefit of the members, the section has organized 4 Continuing Education (CE) classes with relevant topics conducted by well known speakers in applied statistical areas. Significant efforts have been made to distribute the large numbers of section activities across five days and across four morning and afternoon sessions so that sessions on similar topics do not overlap with each other. Attention was paid to ensure that the topics are diverse enough to attract section members with different interests. There are important sessions and good speakers presenting on Sunday afternoon and Thursday morning. I encourage our members to attend these Sunday and Thursday sessions to disprove the myth that these sessions are not well-attended.

The following are the list of the CE classes and Invited sessions organized by Biopharmaceutical section:

Biopharmaceutical Section Sponsored Continuing Education (CE) courses*

Design and Analysis of Crossover Designs by Dallas Johnson	Sat, 28 July 8:30 AM to 5:00 PM
Dose Finding in Drug Development by Naitee Ting et. al.	Sun, 29 July 8:00 AM to 12:00 PM
Analysis of Clinical Trials: Theory and Applications by Christy Chuang-Stein et. al.	Mon, 30 July 8:30 AM to 5:00 PM
Statistical Monitoring of Clinical trials: A Unified Approach by Michael Proschan	Tue, 1 Aug 8:30 AM to 5:00 PM

Biopharmaceutical Section Sponsored Invited Sessions*

Fundamental Statistical Questions underlying Clinical Trials	Sun, 29 July 2:00 PM to 3:50 PM
Statistical Issues in High Dimensional Omics Data and Biomarker Discovery	Mon, 30 July 2:00 PM to 3:50 PM
Collection and Assessment of Safety Data in a New Drug Development Program	Mon, 30 July 10:30 AM to 12:20 PM
Toxicogenomics: From Concept to Regulatory Issues	Tue, 31 July 8:30 AM to 10:20 AM
Statistical Analysis in High Throughput Screening Assays	Wed, 1 Aug 10:30 AM to 12:20 PM
25 Years of Noninferiority Trials	Wed, 1 Aug 2:00 PM to 3:50 PM
Statistical validation of surrogate endpoints	Thu, 2 Aug 8:30 AM to 10:20 AM

* The days and times mentioned are true as of the program on 16 April 2007.

In addition to the above sessions primarily sponsored by our section, the section is co-sponsoring 12 more Invited sessions.

Now that the organization of the sessions, luncheons, and CE classes has been completed, I encourage our section members to participate in these sessions and courses in record numbers to make this year's conference a great success.

Every year, the section presents awards for the Best Contributed paper from the JSM of the previous year. The 2006 Best Contributed paper awards will be presented at the Biopharmaceutical Section Meeting (and Mixer) to which all section members are invited. The 2007 Best Student Paper Awards will also be presented at the same meeting.

I look forward to seeing you in Utah in late July.

Amit Bhattacharyya
Program Chair, 2007
Biopharmaceutical Section of ASA ■

2007 FDA/Industry Workshop, September 17-19, Arlington, VA

Dionne Price, FDA and Matilde Sanchez, Arena Pharmaceuticals

The 11th FDA/Industry Statistics Workshop is scheduled for September 17-19 at the Marriott Crystal Gateway in Arlington, VA. It is co-sponsored by the Biopharmaceutical Section of the American Statistical Association and the Food and Drug Administration Statistical Association. The theme of this year's meeting is "Translating Innovation into Practice through Effective Partnerships." Short courses are scheduled for the first day, September 17th, followed by two days of sessions on the science and statistics associated with the development of new medical products (pharmaceuticals, biologics and devices). The workshop has been very popular since its inception because it is designed specifically to bring together statisticians from industry, academia, and the FDA, and it provides a unique opportunity for open dialogue on issues of mutual interest.

The short courses include Casual Inference and its Applications in Clinical Development (by Don Rubin), Adaptive Designs (by Shein Chow), Statistical Methods for Evaluating Tests and Biomarkers in Medicine (by Margaret Pepe), Data Monitoring Committees in Drug Development (by George Rochester, and George Foley) and Bayesian Analysis (by Brad Carlin). A series of plenary and parallel sessions will discuss a wide variety of important, timely issues, including: Global Harmonization, Adaptive Designs, Clinical Trial Strategies, Noninferiority Trials, Statistical Challenges/Issues in Various Therapeutic Areas, Bayesian Trial Designs, Pharmacogenomics, Clinical Trial Simulation and Modeling, Preclinical Statistics, and Non-randomized Trials.

Based on continued positive feedback from previous attendees, luncheon roundtables, with moderated discussion of a wide variety of topics, will take place on Tuesday, September 18th. The Marriott Crystal Gateway is a premier meeting site located minutes from downtown Washington, DC and Reagan National Airport. The stylish, contemporary hotel offers a full suite of amenities and services, including direct access to the Washington metro rail from the lobby. A block of sleeping rooms has been arranged for workshop attendees. To receive the special workshop rates, attendees must book their reservations by August 24th. Please visit www.amstat.org/meetings/fdaworkshop for the preliminary program, a list of workshop organizers, the registration form, and information on hotel reservations. The FDA/Industry Statistics workshop is a unique forum for statisticians to discuss topics of mutual interest. The organizing committee looks forward to seeing you in September. ■

Let's Hear from You!

If you have any comments or contributions, contact the Editor: Philip Pichotta, phone: 203-882-9321, email: pichottapm@optonline.net; or Past Editor: Richard Caplan, phone: 302-885-5915, email: richard.caplan@astrazeneca.com; or Associate Editor: Thomas Dobbins, phone: 267-305-3090, email: thomas_dobbins@merck.com.

We are looking for volunteers to write articles that will be of interest to our members. Some authorless topics that have been suggested include animal studies and veterinary medicine, bioequivalence in biologics and personalized medicine. If you have been working in an area and would like to suggest a topic or volunteer to write, please send us an email. Non-technical articles related to our work are welcome. One example might be an article about outsourcing statistical programming to Asia (India or China). Perhaps someone could write an article about how to effectively work when the statistical programming is outsourced. How is it different from using a regular CRO? How will our function change?

The *Biopharmaceutical Report* is a publication of the Biopharmaceutical Section of the American Statistical Association.